Sheaves for Data Fusion

N. Sale 15th Feb '19
TDA Reading Group

- Based primarily on Michael Robinson's paper:
  Sheaves are the canonical data structure for sensor integration. (Information Fusion vol 36)

- Main Ideas:
  - present a series of 'axioms' for sensor systems which make it increasingly clear that we're describing a sheaf.
  - describe the core data fusion problem in terms of sheaves.
  - motivate the definition of approximate sections in the context of data fusion.
  - interpret sheaf cohomology as describing obstructions to globally consistent data fusion.

- The Axioms:

  1) "The entities observed by the sensors lie in a set, X".

  $\underbrace{\text{entities}}$

  e.g. targets, sensor data, text documents ...

  2) "All possible attributes of the entities lie in one of some collection of pseudometric spaces

  $$\mathcal{A} = \{ (A_1, d_1), (A_2, d_2), \dots \} ."$$

  e.g. aircraft position, temperature, word count ...

  (note: we can always give some attribute the discrete metric if there's no obvious alternative.)

3) "The collection of all sets of entities whose attributes are related forms a topology $\mathcal{T}$ on X."

(note: If the collection is $\mathcal{U}$, we can let $\mathcal{U}$ be a subbase for our topology.)

This topology is often the geometric realisation of some cell complex in practical situations.

4) "Each sensor assigns, to each open subset of entities, a pseudometric space $(A, d) \in \mathcal{A}$."

We denote this $S : \mathcal{T} \to \mathcal{A}$.

For $U \in \mathcal{T}$, $S(U)$ is called the <u>space of observations</u>.

$a \in \prod_{U \in \mathcal{T}} S(U)$ is called an <u>assignment of S</u>.

$s \in S(X)$ is called a global section.

5) "It is possible to transform sets of observations by reducing the size of sensor domains."

i.e. we have <u>restriction maps</u> for each $U \subseteq V$ in $\mathcal{T}$, $S(U \subseteq V) : S(V) \to S(U)$

s.t. $S(U \subseteq U) = id_{S(U)}$ and $S(U \subseteq W) = S(U \subseteq V) \circ S(V \subseteq W)$

e.g. X a set of pixels, S assigns a collection of pixels to a space recording the pixel values. Then the restriction maps could correspond to cropping the image represented by the pixels.

Note that if Axioms 1-5 are satisfied, then S is a <u>presheaf</u> of pseudometric spaces.

It should now be obvious what Axiom 6 is...

6) "Observations from overlapping sensor domains which agree on the intersection uniquely determine an observation from the union of the domains."

i.e. if $\{U_i\}_{i\in I}$ a collection of sensor domains,

and $a_i \in S(U_i)$ satisfy $\forall i,j \in I$:

$$S(U_i \cap U_j \subseteq U_i)(a_i) = S(U_i \cap U_j \subseteq U_j)(a_j),$$

then $\exists a \in S(\bigcup_{i\in I} U_i)$ s.t. $S(U_i \subseteq \bigcup_{j\in I} U_j)(a) = a_i \quad \forall i \in I$.

Moreover if $\{V_i\}_{i\in I}$ forms an open cover of U, and $a, b \in S(U)$ are s.t. $S(V_i \subseteq U)(a) = S(V_i \subseteq U)(b)$ for each $i \in I$, then $a = b$.

Note that if Axioms 1–6 are satisfied, then S is a <u>sheaf</u> of pseudometric spaces.

We now state a much more restrictive axiom, which while unnecessary for stating the general data fusion problem, is necessary for talking about cohomology later.

7) "Each space of observations $S(U)$ has the structure of a Banach space (complete normed vector space), and each restriction map $S(U \subseteq V)$ is a continuous linear map."

note: quite often we can make non-linear restriction maps linear by instead considering distributions over the original attributes.

We are now ready to state the main problem.

- The Data Fusion Problem:

Giving the space of assignments $\prod_{U \in \mathcal{T}} S(U)$ the pseudometric $D(a,b) := \sup_{U \in \mathcal{T}} d_U(a(U), b(U))$, and noting that a global section $s \in S(X)$ yields an assignment $U \mapsto S(U \subseteq X)(s)$, the problem goes as follows:

"Given assignment $a \in \prod_{U \in \mathcal{T}} S(U)$, find the closest global section: $\operatorname{argmin}_{s \in S(X)} D(a,s) = \operatorname{argmin}_{s \in S(X)} \sup_{U \in \mathcal{T}} d_U(a(U), S(U \subseteq X)(s))$."

prop.) If S is a sheaf of Banach spaces (on a finite $\mathcal{T}$), i.e. Ax 1–7 hold, then this problem always has a unique solution given by a projection.

- Approximate Sections:

We will ignore Axiom 7 for now, and in fact 6 as well.

defn.) Given $\varepsilon \geq 0$, we say that an assignment $s \in \prod_{U \in \mathcal{T}} S(U)$ is an <u>ε-approximate section</u> if

$$d_V(s(V), S(V \subseteq U)(s(U))) \leq \varepsilon \quad \forall V \subseteq U.$$

The minimum value of ε for which s is an ε-approximate section is called the <u>consistency radius</u> of s.

prop.) If S is a presheaf of pseudometric spaces, every global section induces a 0-approximate section. If all the attribute spaces are metric, then this correspondence is a homeomorphism.

prop.) If a is an ε-approximate section of a presheaf S of pseudometric spaces whose restriction maps are K-Lipschitz for some K, then the distance between a and the closest global section is at least $\frac{\varepsilon}{1+K}$, i.e. $\inf_{s \in S(X)} D(a,s) \geq \frac{\varepsilon}{1+K}$.

- Cohomology:

Assume Axioms 6 and 7 again. i.e. consider sheaves of complete, normed vector spaces.

Given $\mathcal{U} = \{U_1, \dots, U_n\}$ a finite open cover for $(X, \mathcal{T})$ and a sheaf S of Banach spaces on $\mathcal{T}$, define the <u>Čech cohomology</u> via

$$C^k(\mathcal{U}; S) := \prod_{i_0 < \cdots < i_k} S(U_{i_0} \cap \cdots \cap U_{i_k})$$

$$d^k(a)_{i_0 < \cdots < i_{k+1}} := \sum_{j=0}^{k+1} (-1)^j S(U_{i_0} \cap \cdots \cap U_{i_{k+1}} \subseteq U_{i_0} \cap \cdots \hat{U}_{i_j} \cdots \cap U_{i_{k+1}})(a_{i_0 < \cdots \hat{i_j} \cdots < i_{k+1}})$$

$$H^k(\mathcal{U}; S) := \frac{\ker d^k}{\operatorname{im} d^{k-1}}.$$

Therefore note that when $\mathcal{T}$ is finite, we can talk about $H^k(\mathcal{T}; S)$. Moreover, we can often make this easily computable via the following

Leray Theorem) Suppose S a sheaf on $\mathcal{T}$ and $\mathcal{U} = \{U_i\}$ is a collection of open sets s.t. for each intersection $U = U_{i_0} \cap \cdots \cap U_{i_k}$ of a finite number of elements in $\mathcal{U}$, we have $H^k(U \cap \mathcal{T}; S) = 0 \ \forall k > 0$, then $H^p(\mathcal{U}; S) = H^p(\mathcal{T}; S) \ \forall p$.

prop.) If S is a sheaf of vector spaces on a finite topology $\mathcal{T}$, then $H^0(\mathcal{T}; S) = \ker d^0$ is the space of assignments corresponding to global sections of S.

proof: Immediate from writing

$$d^0(a)_{i_0 < i_1} = S(S_{i_0} \cap S_{i_1} \subseteq S_{i_1})(a_{i_1}) - S(S_{i_0} \cap S_{i_1} \subseteq S_{i_0})(a_{i_0})$$

and axiom 6.

Note, a nontrivial element of $H^k(\mathcal{U}; S)$ consists of a collection of observations on the k-way intersection of sensor domains in $\mathcal{U}$ that are consistent on further restriction (i.e. $\in \ker d^k$), but which don't arise from any $(k-1)$-way intersections ($\notin \operatorname{im} d^{k-1}$).

These are classes of self-consistent data that can't be extended to global sections.

They typically arise when the underlying model allows for some sort of inconsistency, perhaps due to the assumptions of the model.

- An Example:

Consider the following sheaf model of some weather sensors:

$$\begin{tikzcd}
S(\text{News}) \arrow[rr, "\neg id"] \arrow[dd, "id"'] & & S(\text{Raining?}) \arrow[dl, "\begin{pmatrix}0&0\\0&1\end{pmatrix}"] & & S(\text{Humidity}) \arrow[ll, "id"'] \arrow[ddl, "0"] \\
& S(\text{cloudy?}) & & & \\
S(\text{Sunny?}) \arrow[ur, "\neg id"'] & & & S(\text{Raining literal cats and dogs}) \arrow[ull, "\begin{pmatrix}0&0\\0&1\end{pmatrix}"'] & \\
& S(\text{Camera}) \arrow[ul, "id"] \arrow[urr, "0"'] & & &
\end{tikzcd}$$

(Corresponding to the topology on .)

where we interpret boolean values as lying in $\mathbb{F} \oplus \mathbb{F}$, with true $= \begin{pmatrix}0\\1\end{pmatrix}$ and false $= \begin{pmatrix}1\\0\end{pmatrix}$, $id = \begin{pmatrix}1&0\\0&1\end{pmatrix}$, $\neg id = \begin{pmatrix}0&1\\1&0\end{pmatrix}$.

(So note that $\begin{pmatrix}0&0\\0&1\end{pmatrix}$ means that if it's raining (cats + dogs), then it must also be cloudy, but otherwise we don't know).

Then we can calculate (I have a Haskell script if you're interested)

$$H^0(\mathcal{T}; S) = \langle (\text{News, Camera} = \text{true}, \text{Humidity} = \text{false}), (\text{News, Camera} = \text{false}, \text{Humidity} = \text{true}) \rangle$$

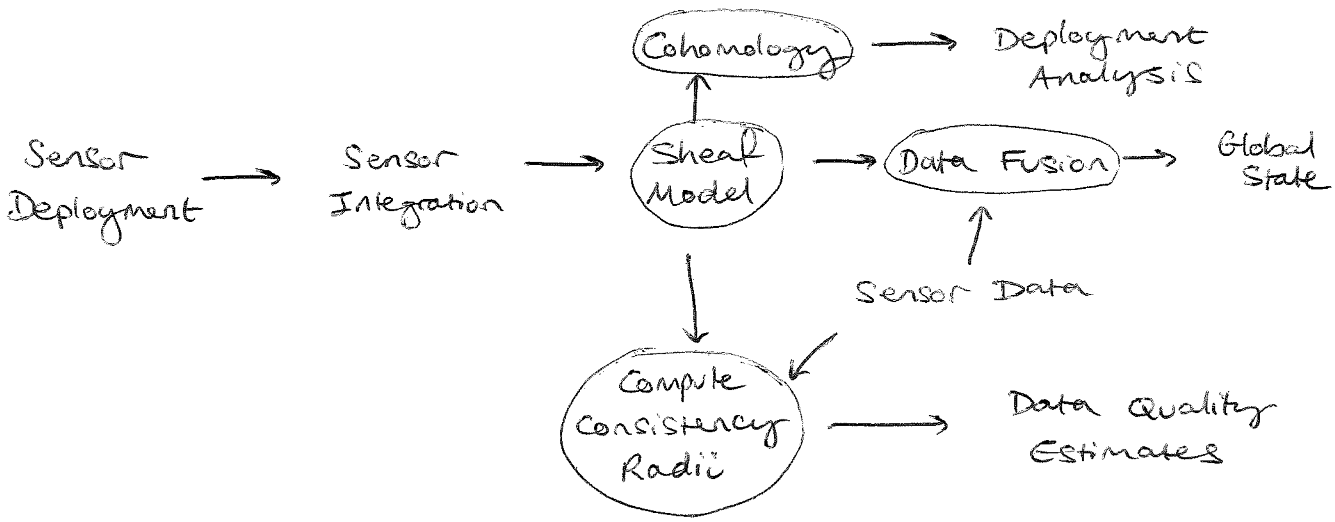
(it can't be sunny and raining at the same time).

$$H^1(\mathcal{T}; S) = \langle (\text{Raining cats + dogs} = \text{false}), (\text{Raining cats + dogs} = \text{true}) \rangle.$$

(we could consistently assume ~~was~~ it is, or isn't raining literal cats and dogs, but this data can't have come from any global section.)

- Summary:

  I will give a summary in the form of the framework Robinson envisions:



we've looked at the circled components.